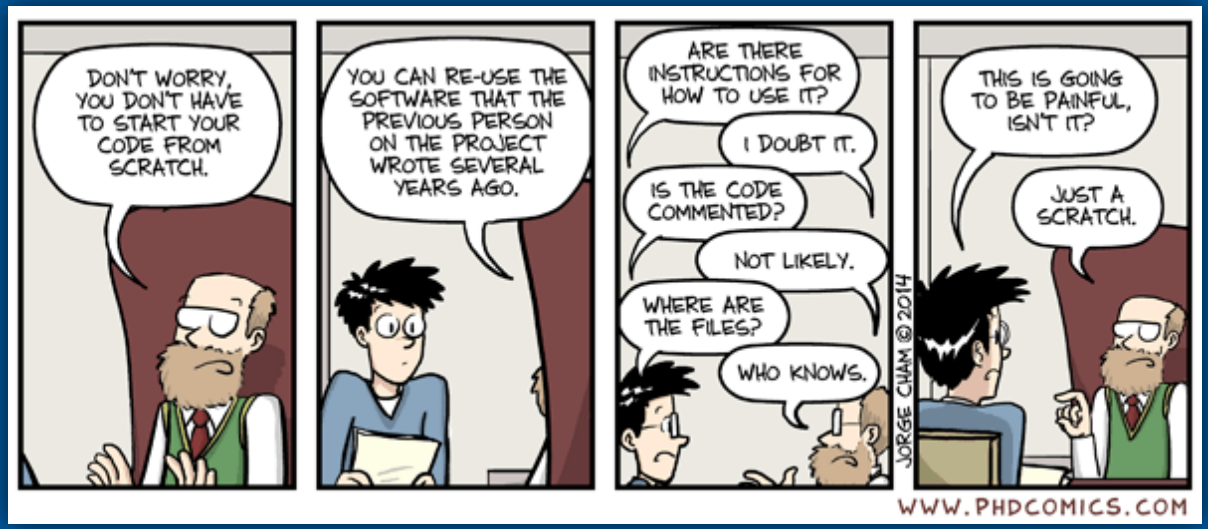
# Integrierte Entwicklungs- und Publikationsumgebung für Forschungssoftware und Daten am Helmholtz-Zentrum Dresden-Rossendorf (HZDR)

Tobias Frust
Guido Juckeland
Uwe Konrad




RODARE
ROSSENDORF DATA REPOSITORY



HZDR
HELMHOLTZ ZENTRUM DRESDEN ROSSENDORF

# HZDR – Facts and Figures

- Member of the Helmholtz Association

- **Foundation** 01.01.1992 (e.V.)
  Forschungszentrum Rossendorf

- **Employees** **~ 1.200**
  including about 500 scientists
  + 150 PHD students;
  as well as employees and guest scientists from more than 50 countries

- **Research Sites** **Dresden**
  Helmholtz-Institut **Freiberg**
  Forschungsstelle **Leipzig**
  HIBEF-Station am XFEL **Schenefeld**
  Rossendorf Beamline an der
  ESRF in **Grenoble**

Bilder: Killig, DESY, ESRF/Ginter

# Large Research Infrastructures

## ELBE - Center for High-Power Radiation Sources

- Electron accelerator ELBE feeds free-electron lasers FELBE & THz source TELBE;
- generates positrons, protons and neutrons as well as X-ray and gamma radiation;
- plus high-intensity lasers (1 Petawatt) **DRACO** and **PENELOPE** (under construction)

## Dresden High Magnetic Field Laboratory (HLD)

- Producing Europe's highest pulsed magnetic fields for materials research

Bilder: Bierstedt, Killig (2 x)

## Ion Beam Center (IBC)

- Nanoscale surface analysis and modification

# Research data and software – Where do you store it?

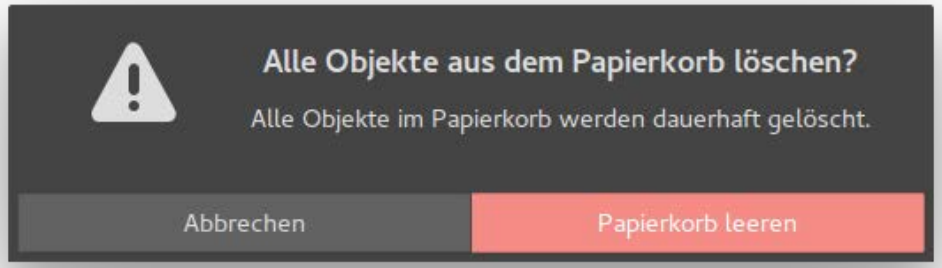
?

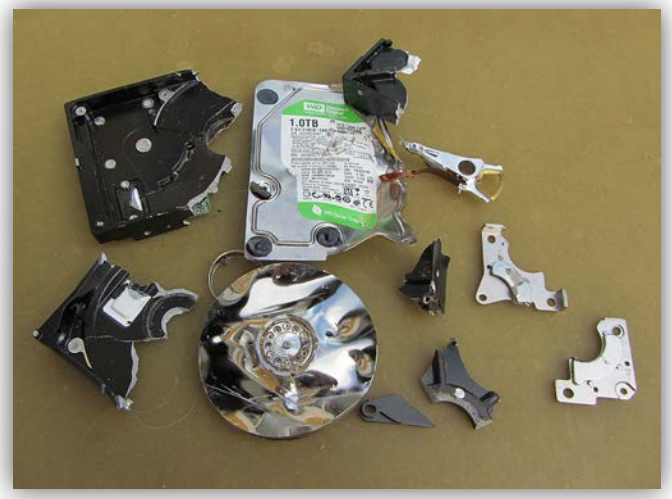
RODARE

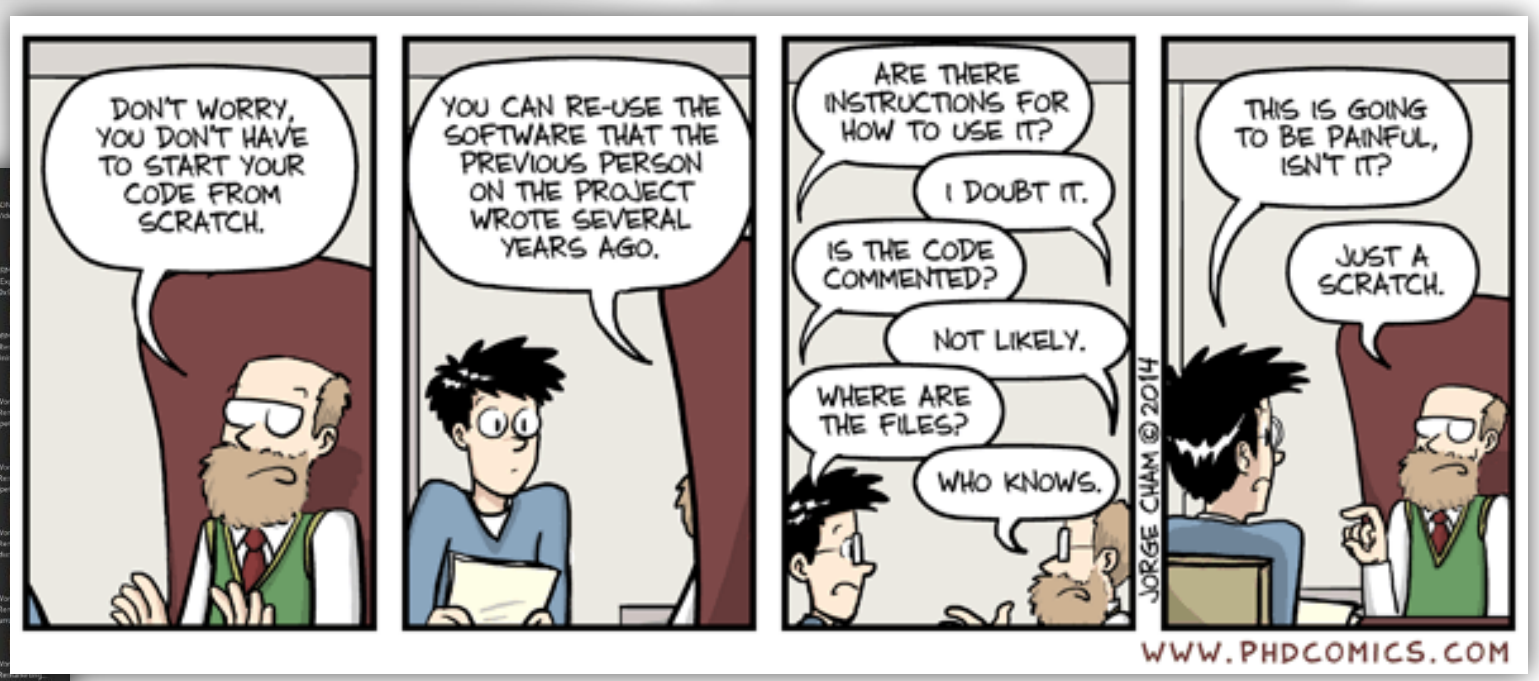
DRESDEN concept

HZDR

# Research data and software – Is it safe?

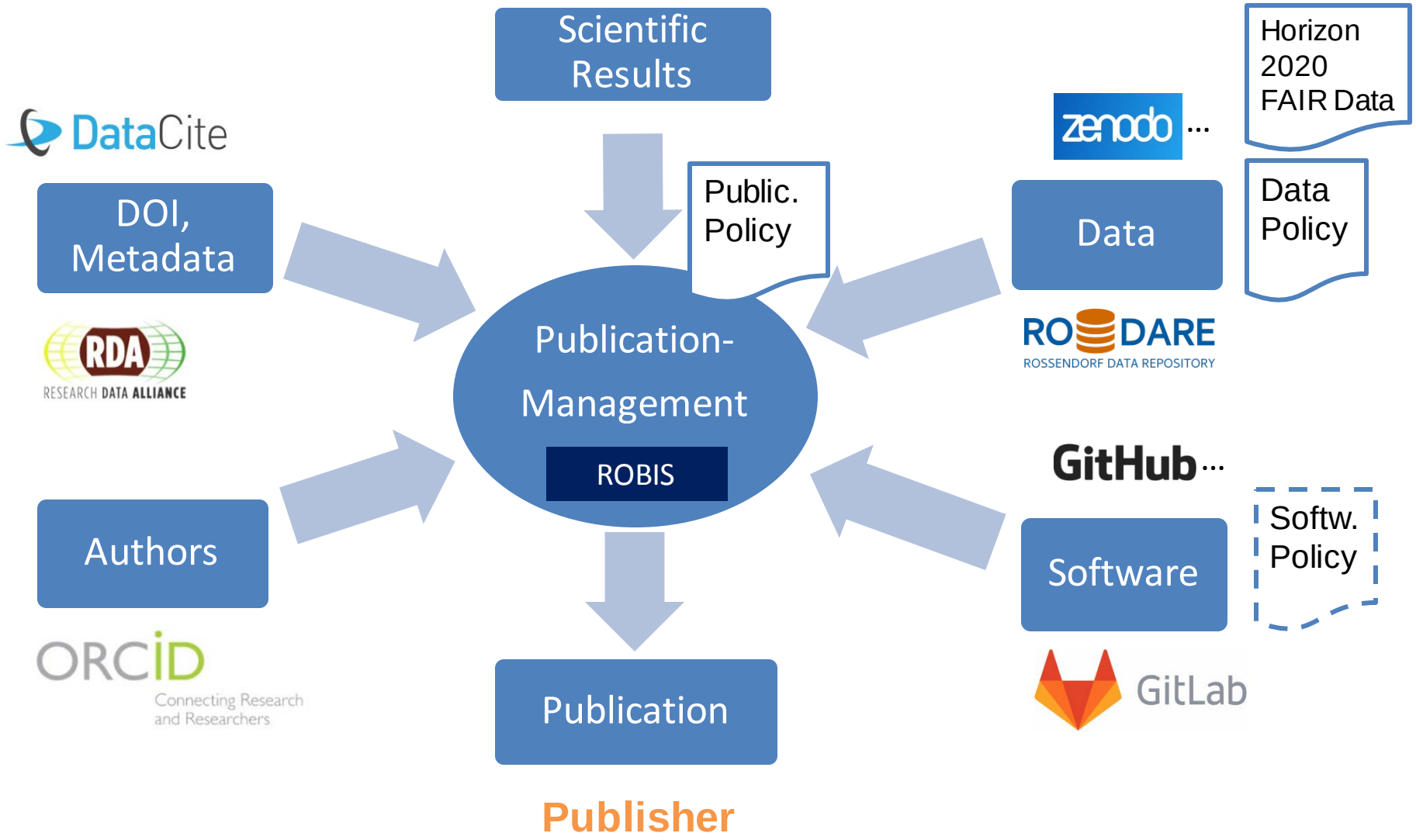




RODARE

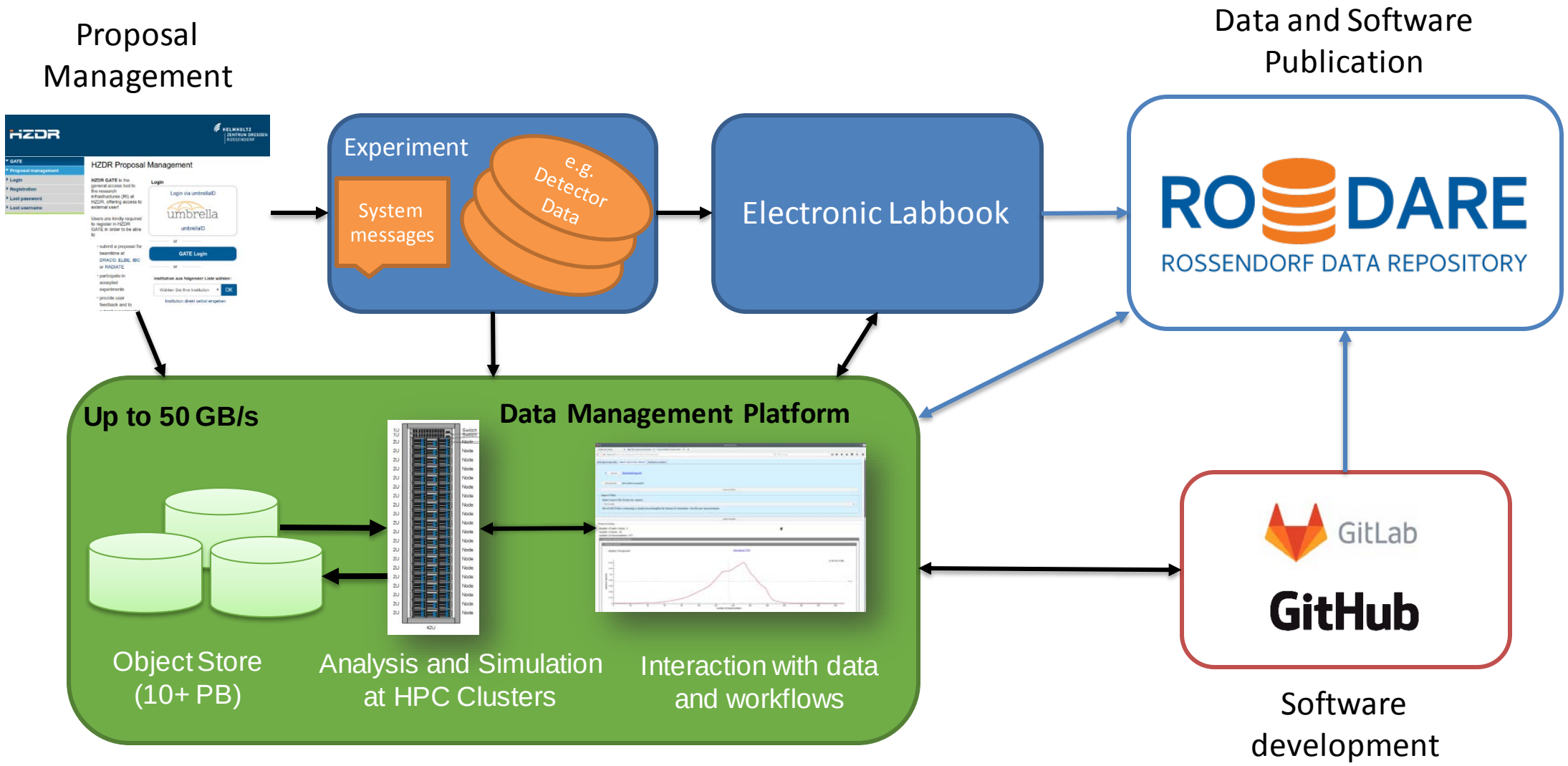
DRESDEN concept

HZDR

# Publication Components



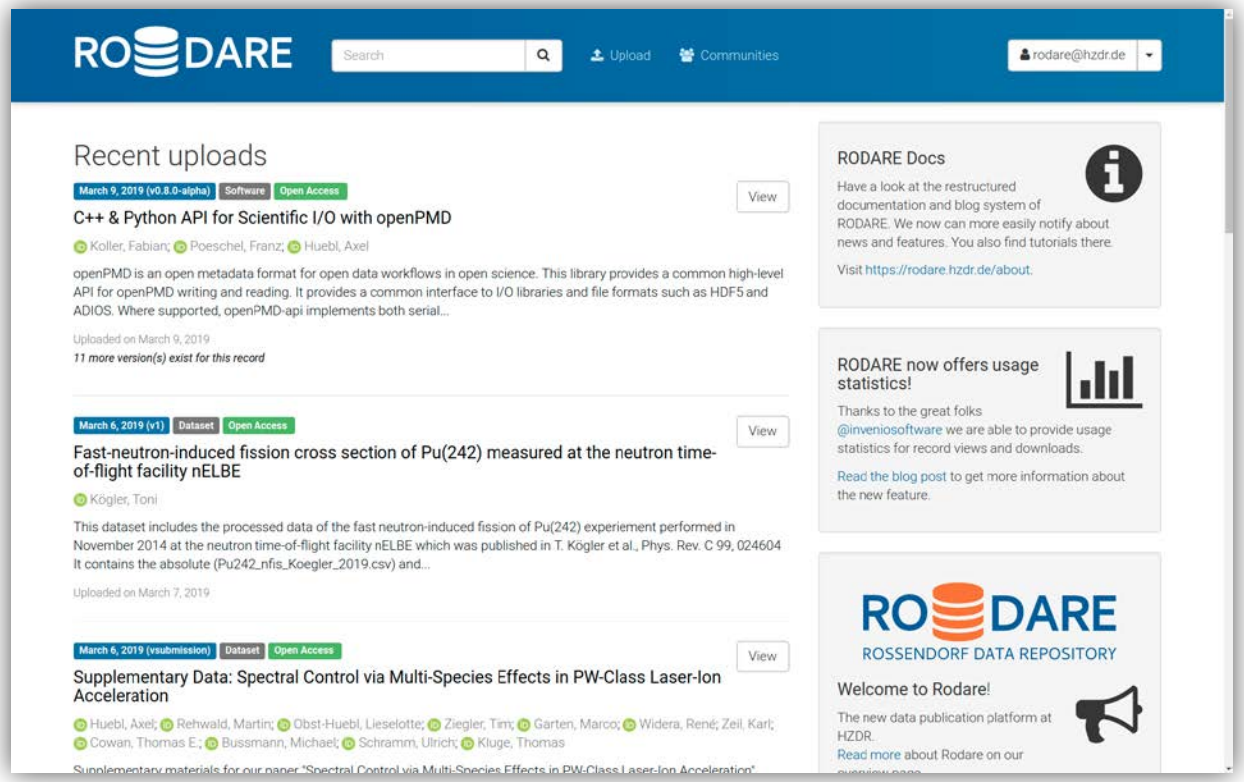
# Services for the whole lifecycle of data and software



# Research data and software – How to do it right?

RODARE
ROSSENDORF DATA REPOSITORY

https://rodare.hzdr.de




**Why not simply Zenodo?**

- Filesize limit 50 GiB; Dataset size limit 100 GiB
- → Not enough for the datasets we need to publish
- → Even if we had infinite space, bandwidth limitation
- Custom integrations for local experiments are planned

RODARE

DRESDEN concept

HZDR

# Challenges we are facing at HZDR

- Many users don’t see the need for open science yet.

→ *Conviction process is required - takes time. Journal requirements regarding open data/software often help.*

- Publishing research data or software is often seen as only an additional burden.

→ *Reduce the barrier by providing good integrations and establish obligatory guidelines.*

- Knowledge about Open Science and “How to do it right” is often not available.

→ *Training and support is required.*

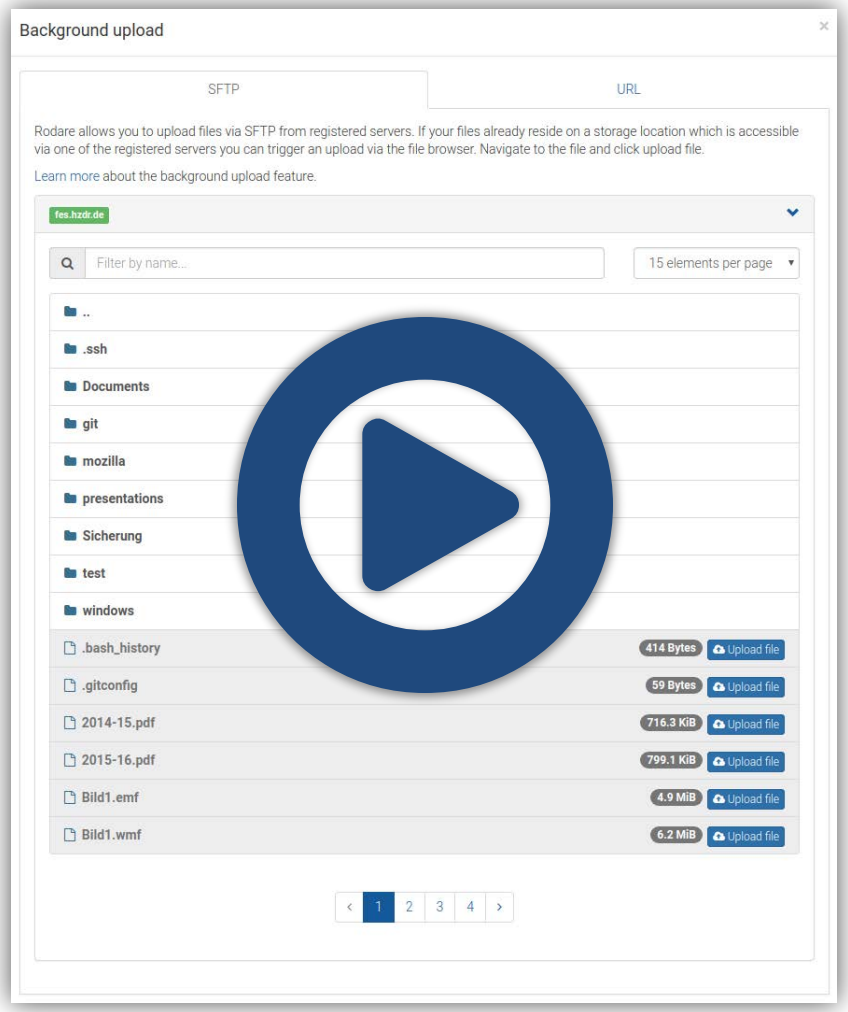
Photo by Mārtiņš Zemlickis on Unsplash

# Integrations – Background Upload

- Users need to be able to upload huge files (> 100 GiB) – web upload is error prone.
- Many files already reside on central storage servers – but under the control of the user.

→ Background upload via SFTP

→ Completely controllable via Web UI or Rest API



# Integrations - GitLab

- Publishing software releases should be as easy and as automated as possible.
- GitHub integration for Invenio already exists.
- A lot of institutions self-host a GitLab instance or use Gitlab.com, and so is HZDR.

→ Integration for private GitLab instances or GitLab.com is a popular request.

→ **GitLab-Integration for Invenio:** https://gitlab.hzdr.de/rodare/invenio-gitlab

→ Primary developed for Rodare; will hopefully be integrated into Zenodo, too.


RODARE

DRESDEN concept HZDR

# Policies

- **HZDR Data Policy** (since May, 2018). Instructions for:
  - Responsibilities
  - Data Management Plan (DMP)
  - Embargo period
  - Licensing of research data
- **Software Policy**
  - Under construction in collaboration with the Helmholtz Task Group Open Science
  - First draft of a template is now discussed with other responsibilities (e.g. Technology Transfer, Legal department, …)

→ **Data Librarian:** Provide support and training for scientists

# Invenio RDM [1] – The Future of RODARE

The project has an ambitious one year schedule to build an RDM solution based on Zenodo. Invenio RDM will deliver:

- Invenio RDM - A research data management platform based on Zenodo and Invenio v3.
- A community of public and private institutions to sustain Invenio RDM.
- Minimum two existing repositories migrated to Invenio RDM, with Zenodo being one of them.


[1] https://invenio-software.org/blog/2019-04-29-rdm/

# Invenio RDM [1] – The Future of RODARE

- → **Core repository:** extensible metadata model based DataCite metadata schema with support for handling millions of records and peta bytes of data
- → **Packaging and distribution:** require minimal experience in installing, operating and administering the platform
- → **Customization and extendability:** easy extensibility and customisability to adapt to each particular institution


RODARE [1] https://invenio-software.org/blog/2019-04-29-rdm/

# Conclusion

- Publication of research data and software requires **integrated services**.
- Metadata and documentation should be captured **continouosly and as soon as possible**.
- Metadata should be collected **as automated as possible**.
- Users **need adapted and individual** support and training.

For the service provider:

- Installation, administration and deployment should be as **easy** as possible.
- Easy **extensibility and customizability.**
- Use **widely accepted** standards, e.g. for metadata formats.

# Thank you!

What do you think?

Are we doing data and software publication right?

RODARE

RODARE
ROSSENDORF DATA REPOSITORY

Powered by
INVENIO

DRESDEN concept

HZDR